

A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Douglas A. Chrissan *et al.*  
Docket: 8X8S.239PA  
Title: VARYING PULSE AMPLITUDE MULTI-PULSE ANALYSIS SPEECH  
PROCESSOR AND METOHD

CERTIFICATE UNDER 37 CFR 1.10

Express Mail' mailing label number: EL449833116US

Date of Deposit: September 8, 1999

I hereby certify that this paper or fee is being deposited with the United States Postal Service 'Express Mail Post Office To Addressee' service under 37 CFR 1.10 and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

By: 

Name: Kelly J. Ledin

BOX PATENT APPLICATION  
Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

We are transmitting herewith the attached:

- ☒ Transmittal sheet containing Certificate under 37 CFR 1.10.
- ☒ Patent Application: Pages Numbered 1-23; 32 claims; Abstract 1 pgs.
- ☒ 2 sheets of informal drawings
- ☒ An unexecuted Declaration
- ☒ PAYMENT OF THE FILING FEE IS BEING DEFERRED
- ☒ 1 return postcard

The fees are calculated as follows:

Basic Filing Fee = \$760.00

Fee for total number of claims in excess of 20 =  $\$18 * (32 - 20) = \$216.00$

Fee for total number of independent claims in excess of 3 =  $\$78 * (10 - 3) = \$546.00$

Total fee due = \$1,522.00

CRAWFORD PLLC  
333 Washington Avenue North, Suite 5000  
Minneapolis, MN 55401  
(612) 349-2700

By: 

Name: Robert J. Crawford

Reg. No.: 32, 122



Patent Application for:

**VARYING PULSE AMPLITUDE MULTI-PULSE ANALYSIS**

**SPEECH PROCESSOR AND METHOD**

Invented By:

**Douglas A. Chrissan**

a citizen of the U.S.A. and

**Rajarathinam G. Subramanian**

a citizen of India, and

each an employee of the Assignee:

8x8, Inc.

of

Santa Clara, California

668060-1472660

**VARYING PULSE AMPLITUDE MULTI-PULSE ANALYSIS**  
**SPEECH PROCESSOR AND METHOD**

5

**Field of the Invention**

The present invention relates generally to speech signal processing and, more particularly, to multi-pulse speech analysis and synthesis systems.

10

**Background of the Invention**

Speech signal processing is well known in the art and is often utilized to compress an incoming speech signal for applications such as storage and transmission. The speech signal processing typically involves dividing the incoming speech signals into frames and then analyzing each frame to determine its representative components. The representative components are then stored or transmitted.

15

A frame analyzer is often used to determine the short-term and long-term characteristics of the speech signal. The frame analyzer can also determine one or both of the short- and long-term components, or contributions, of the speech signal. As an example, linear prediction coefficient (LPC) analysis provides the short-term characteristics and contribution, and pitch analysis and prediction provides the long-term characteristics as well as the long-term contribution.

20

Typically, one, both or neither of the long- and short-term predictor contributions are subtracted from the input frame, leaving a target vector whose shape has to be characterized. Such a characterization can be produced with multi-pulse analysis (MPA) which is described in detail in section 6.4.2 of the book *Digital Speech Processing, Synthesis and Recognition* by Sadaoki Furi, Marcel Dekker, Inc., New York, N.Y. 1989, incorporated herein by reference.

25

Conventionally, MPA involves a target vector that is formed of a multiplicity of samples. The target vector is modeled by a plurality of pulses of equal amplitude varying location and

varying sign (positive and negative). To select each pulse, a pulse is placed at each sample location and the effect of the pulse, defined by passing the pulse through a filter defined by the LPC coefficients, is determined. The pulse which provides the filter output that most closely matches the target vector is selected and its effect is removed from the target vector, thereby  
5 generating a new target vector. The process continues until a predetermined number of pulses have been found. For storage or transmission purposes, the result of the MPA analysis is a collection of pulse locations, pulse signs (positive or negative), and a quantized value of the pulse amplitude.

The MPA output typically specifies the resulting pulse locations, but not the order in  
10 which they were chosen. It also specifies only one gain parameter, so the decoder must reconstruct the pulse sequence using equal amplitudes for all the pulses. In addition, the MPA analysis itself is sub-optimal, from a maximum-likelihood standpoint, with respect to determining the best possible pulse sequence to match the target.

Accordingly, there is need for a speech processor and method that improves the  
15 performance of the MPA process and the perceptual quality of the reconstructed speech and that overcomes the above-mentioned deficiencies of the prior art.

### **Summary**

According to certain embodiments, the present invention provides a speech processing  
20 method and arrangement including a process applicable for use in connection with the ITU G.723.1 speech encoding recommendation. Certain embodiments of the invention are applicable to multipulse maximum likelihood quantization coding systems and processes.

Particular embodiments involve method and structure approaches directed to speech  
processing systems in which a signal processor arrangement analyzes an input speech signal and,  
25 in response, generates the short-term characteristics of the input speech signal and a target vector. One such approach involves: generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences

having a different average amplitude value; and outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

Another particular application of the present invention involves a speech processing method and arrangement that utilizes pulse sequences of varying amplitude pulses in one or both of the MPA unit and the decoder. In particular embodiments, pulse sequences of varying amplitude pulses are used in each of the MPA unit and the decoder; the MPA unit and not the decoder; and the decoder and not the MPA unit. The digitally compressed representative signal need not contain additional information about the variation of the pulse amplitudes or about the order in which the pulses were chosen.

In these particular applications, the pulse amplitude variation within a given sequence is typically small relative to the average amplitude of the sequence. A typical ratio is 20-30 percent.

One important aspect of the present invention is directed to the performance of the MPA process and the perceptual quality of the reconstructed speech. Consistent with this aspect of the present invention, another particular example embodiment involves a speech processing system that includes a short-term analyzer, a target vector generator and a multi-pulse analysis unit. The system optionally includes a long-term analyzer, and the MPA unit can use a maximum-likelihood criterion for evaluating the error of a given pulse sequence. The target vector is generated from the input speech signal or a perceptually modified version of the input speech signal, and the MPA unit operates on at least the target vector and the short-term characteristics determined by the short-term analyzer.

In another particular example embodiment of the present invention, the MPA varies the amplitudes of the pulses in each pulse sequence when choosing the pulse locations within a given pulse sequence, and utilizes equal amplitude pulses when determining the best pulse sequence based on the given error criterion. In another embodiment of the present invention, the encoder varies the amplitudes of the pulses in each pulse sequence when determining the best pulse sequence based on the given error criterion, but the decoder does not have knowledge of these pulse amplitude variations. In a third embodiment of the present invention, both the encoder and

the decoder have knowledge of the variation of the pulse amplitudes in a given pulse sequence. The encoder takes these amplitude variations into account when choosing the pulse locations within a given pulse sequence and/or when choosing the best pulse sequence based on the given error criterion. The encoder and decoder may utilize one or both of: a predetermined pulse modification function, and a pulse modification function derived from parameters or signals known by both the encoder and decoder. Example signals known by both the encoder and decoder are: the LPC parameters (short-term characteristics), the long-term pitch parameters (long-term characteristics), and the previous excitation signal.

Another embodiment of the present invention uses pulse-train sequences instead of pulse sequences. Each pulse train in a pulse-train sequence consists of equal amplitude, equal sign, equally spaced pulses, and the different pulse trains have varying amplitudes.

Another embodiment of the present invention uses pulse-train sequences with each pulse train in a pulse-train sequence consisting of variable amplitude, variable sign, and equally spaced pulses. Further, the different pulse trains have varying average amplitudes.

In other embodiments, the above embodiments are combined in one of various ways for a given system and application. In one system, for instance, both a varying-amplitude multi-pulse pulse sequence analysis and a varying-amplitude multi-pulse pulse train analysis are performed and the one resulting in the closest match to the target vector is chosen as the MPA unit's output signal.

The above summary of the invention is not intended to describe each disclosed embodiment of the present invention. An overview of other example aspects and implementations will be recognizable from the figures and of the detailed description that follows.

### **Brief Description of the Drawings**

The present invention will be understood and appreciated more fully from the following detailed description taken in conjunction with the drawings in which:

FIG. 1 is a block diagram illustrating a speech processing system, according an example embodiment of the present invention; and

FIG. 2 is a flow chart illustrating speech processing, according an example approach that is consistent with the present invention.

5 While the invention is amenable to various modifications and alternative forms, specifics thereof have been shown by way of example in the drawings and will be described in detail. It should be understood, however, that the intention is not to limit the invention to the particular embodiments described. On the contrary, the intention is to cover all modifications, equivalents, and alternatives falling within the spirit and scope of the invention as defined by the appended  
10 claims.

### **Detailed Description**

The present invention is generally applicable to speech processing arrangements involving multi-pulse signal representation where accurate signal representation is important to system operation. The present invention has been found to be particularly advantageous for systems of this type when implemented in compliance with conventional speech encoding systems, such as those intending to be compliant or compatible with the ITU G.723.1 and other speech encoding recommendations involving multipulse coding arrangements and methods.

15 A particular example application is a video and speech encoding/decoding system such as used for videoconferencing. Such a system is described in connection with U.S. Patent Application No. 09/005,053, filed on January 9, 1998 (Docket No. 11611.51US01), which is incorporated herein by reference. The example video-control units and video-processing circuits illustrated and described therein employ a multiple-processor structure including a digital signal processor ("DSP") and a RISC processor. The DSP is arranged to handle specialized tasks such  
20 as compression and decompression of video and speech information, and the RISC processor is arranged to process most other functions. Alternatively, this example speech-processing embodiment is implemented using a dedicated DSP.

An appreciation of the various advantages and aspects of the invention can be realized using such an example videoconferencing application. For the purpose of conveying these various advantages and aspects, FIG. 1 and its related discussion illustrate various example embodiments of the present invention in the context of a speech-processing arrangement and as may be used in a videoconferencing system such as described above.

Reference is now made to FIG. 1, which generally illustrates an example embodiment of the present invention as applied to a speech-processing application. The depicted speech processing system includes various functional blocks, including a short-term prediction analyzer 10, a long-term prediction analyzer 12, a target vector generator 13 and a multi-pulse analysis (MPA) unit 14. The functions of the short-term prediction analyzer 10, the long-term prediction analyzer 12, and the target vector generator 13 can be implemented in any of a number of ways to process input frames of a speech signal formed of a multiplicity of digitized speech samples.

In one example, input speech is in the form of 240 speech samples per frame, each frame is separated into a plurality of four subframes, and each subframe is sixty samples long. The input frame can be a frame of an original speech signal or of a processed version thereof. The short-term prediction analyzer 10 receives the input frame and produces on signal line 17, the short-term characteristics of the input frame. In one specific embodiment, short-term prediction analyzer 10 performs linear prediction analysis to produce linear prediction coefficients (LPCs) that characterize each input frame, and with each subframe being processed one at a time.

The long-term predictor analyzer 12 also operates on the input frame received on line 16. The long-term analyzer 12 analyzes a plurality of subframes of the input frame to determine the pitch value of the speech within each subframe, where the pitch value can be defined as the number of samples after which the speech signal approximately repeats itself. In many applications, pitch values typically range between 20 and 146, where 20 indicates a high-pitched voice and 146 indicates a low-pitched voice.

Once the long-term analyzer 12 determines the pitch value, the pitch value is utilized to determine the long-term prediction information for the subframe, provided on the signal line 18.



The target vector generator 13 outputs a target vector for processing by the MPA unit 14 in response to the output signals of the long-term analyzer 12 and of the short-term prediction analyzer 10 and in response to the input frame, via a delay 19. Using these signals, target vector generator 13 generates the target vector from one or more subframes of the input frame. Various aspects of the long-term and short-term information can be utilized, if desired, or they can be ignored. The delay 19 is used to delay the input frame so that it arrives at the target vector generator 13 so as to correspond to the respective outputs of the analyzers 10 and 12.

As indicated in FIG. 1, the MPA unit 14 receives as inputs a short-term impulse response (IR) and a target vector along signal lines 17 and 26, respectively. Using the example system front-end feeding the MPA unit 14, the short-term impulse response is, or produced as part of, the short-term characteristics produced by the short-term prediction analyzer 10, and the target vector is received from the target vector generator 13. In another front-end embodiment, the short-term impulse response (IR) is received from a short-term analyzer and a target vector is received from another type of target vector generator.

The MPA unit 14 of FIG. 1 includes various functional blocks. These blocks are a signal correlator and gain-range determiner (SC/GRD) 22, a pulse amplitude selector 24, a pulse sequence determiner 25, a target vector matcher 28, and an optional encoding unit 30. These blocks process the short-term impulse response and the target vector, according to the present invention, to identify and encode one of a number of the pulse sequence candidates that best matches to the target vector. Such an encoded pulse sequence and its parameters are shown as being output of the MPA unit 14.

Using the received short-term impulse response and the target vector, the SC/GRD 22 calculates the autocorrelation of the impulse response and the cross-correlation between the impulse response and the target vector. This calculation can be accomplished, for example, as in the above-referenced ITU G.723.1 speech processing recommendation.

These two correlation signals are used by the SC/GRD 22 to determine an initial pulse gain and pulse location. In one specific implementation, the initial pulse gain and pulse location are determined as presented in the above-referenced ITU G.723.1 speech processing

recommendation. The amplitude of a given pulse sequence to be searched is typically referred to as the quantized gain of the pulse sequence, and in many embodiments a range of gains is searched. In the example embodiment of FIG. 1, the range of gains searched, the pulse gain and pulse location of the current pulse sequence are output to the pulse amplitude selector 24. The two correlation signals calculated by the SC/GRD 22 are also output to the pulse sequence determiner 25, as depicted at signal line 27.

The pulse amplitude selector 24 receives the gain range and moves through the gain values within the gain range that was obtained from the SC/GRD 22. The pulse amplitude selector 24 then outputs the pulse amplitude of the current pulse sequence, depicted at signal line 32. This current pulse sequence, as provided at signal line 32, is a current gain level for which a sequence of pulses is to be determined.

The pulse sequence determiner 25 receives the two correlation signals from the SC/GRD 22 (signal line 27) and the pulse amplitude of the current pulse sequence from the pulse amplitude selector 24, and performs a multi-pulse analysis to determine the signs and locations of the pulses in the pulse sequence. The current pulse sequence on output line 34 is analyzed by the target vector matcher 28, which compares the fit of the current pulse sequence to the target vector with the fit of previously analyzed pulse sequences based on the given error criterion. For each gain value, the target vector matcher 28 determines the quality of the match, saving the match (gain index and pulse sequence) only if it provides a smaller value for the criterion than the value associated with previous matches. If the present pulse sequence provides a better match to the target vector than the value associated with any of the previous sequences, its pulse signs and locations and gain are stored. After all candidate pulse sequences are determined and matched to the target vector, the one resulting in the best match to the target vector is output to the encoder on line 38. Since there are a range of gain levels, the matcher 28 returns control to the gain level selector 24 to select the next gain level. This return of control is indicated by arrow 36.

For the target vector matcher 28, the given error criterion can be implemented, for example, as described in connection with a maximum likelihood criterion, or a minimum mean

squared error criterion. For further information pertaining to such processing (and for other related speech processing information), reference may be made to the book entitled, "Probabilistic Methods of Signal and System Analysis," 3rd Edition, George R. Cooper and Clare D. McGillen, Oxford Univ. Press, 1999, and the book entitled "Digital Speech Coding for Low Bit Rate Communication Systems," by A.M. Kunooz, John Wiley & Sons, Ltd., West Sussex, England, 1994. As alternatives to such a maximum likelihood criterion, a perceptual quality criterion implemented with empirical testing may be used.

The best match provided by the target vector matcher 28 is then encoded by the optional encoding unit 30 and its parameters are presented at the output of the MPA unit 14. The pulse sequence is typically represented as a series of positive and negative pulses having the current gain level. Optional encoder 30 encodes the output pulse sequence and gain index for storage or transmission.

The SC/GRD 22 of FIG. 1 can be implemented using various approaches. One approach is an embodiment described in U.S. Patent No. 5,568,588. In this patent, a gain range determination is made to determine an amplitude of the first pulse and then a range of quantized gain levels around the absolute value of the determined amplitude based on a fixed number of steps for moving through the range of quantized gain levels. This relates to the approach of the ITU G.723.1 speech coding recommendation, which is based on the number of steps being fixed at four, for moving through a set range of quantized gain levels.

Another approach is illustrated and described in the above-referenced U.S. Patent Application No. 09/086,434 (8X8S.200PA). The step size (referred to as MLQ\_STEPS) is provided by the MPA unit 14. As applied to the example embodiment of FIG. 1, the gain range determination is a function of the first pulse output of a pulse location determination, an initial quantized gain level, and a set of selected quantized gain levels to be searched as a function of the initial quantized gain level. Both MLQ STEPS and the range of unquantized gain levels searched are a function of the initial quantized gain level, or equivalently, the absolute value of the determined amplitude.

FIG. 2 is a flow chart showing an example manner in which the system of FIG. 1 can be implemented according to the present invention. The example flow begins at block 50, which corresponds to the SC/GRD 22 of FIG. 1. Block 50 determines the IR autocorrelation, the target vector impulse response (TV-IR) cross correlation, and the gain range, as described above and in connection with the ITU G.723.1 speech coding recommendation.

At block 51, the pulse amplitude is selected as described above in connection with the pulse amplitude selector 24 of FIG. 1.

From block 51, flow proceeds to block 52 where the pulse amplitude is modified as a function of pulse amplitude modification parameters. In particular embodiments, these pulse amplitude modification parameters are provided to improve and/or change the perceptual quality of the reconstructed speech by various experimental or methodical approaches. An example of an experimental approach is the result of empirically testing and, therefrom, defining these pulse amplitude modification parameters. An example of a methodical approach involves establishing these pulse amplitude modification parameters as exponentially-based function, as described more fully below. Block 52 modifies the pulse amplitude of each pulse in a given sequence during the location search phase of the MPA unit 14.

From block 52, flow proceeds to block 53 where the pulse location is determined and is optionally modified to apply a selection bias that varies with location in the analysis frame. The pulse location determination operation at block 53 uses pulses of varying amplitude within a given pulse sequence. This selection bias can optionally change on a frame-by-frame basis. Further, at block 53, the pulse contributions, as provided in connection with block 52, are removed. This removal can be readily accomplished by, for example, by subtracting each pulse's contribution to the reconstructed signal from the target vector. For a given sequence, the operations of block 53 are executed once for each pulse.

Block 54 reflects the determination of whether to return to block 52 if there are additional pulses to choose or, if there are not additional pulses to choose, to proceed to block 55. Like the pulse location determination operation at block 53, the pulse sequence reconstruction at block 55 uses pulses of varying amplitude within a given pulse sequence.

The operation at blocks 53 and 54 can be implemented as part of the pulse sequence determiner 25 of FIG. 1.

At block 55, the pulse sequence is reconstructed using the digitally coded information received for each pulse (including the pulse sequence's reference, or "central", pulse amplitude, and the location and sign of each pulse) to execute the coder's (encoder and/or decoder) predetermined reconstruction implementation of the sequence around the central pulse amplitude. With the exception of the introduced pulse amplitude variability issues, such reconstruction is conventional and may be implemented, for example, as characterized in the above-mentioned ITU recommendation. The skilled artisan will appreciate that: in various implementations, the term "central pulse amplitude" can refer to the average amplitude value, the median amplitude value, or any centrally-located value within the range of the pulse amplitudes in a given sequence; that the degree of variability introduced in the pulse amplitude can be limited by whether the coder's (encoder or decoder) reconstruction implementation is manufacturer-compatible with the communicatively-coupled coder's (decoder or encoder) reconstruction implementation; and that the coder's (encoder or decoder) reconstruction implementation can be negotiated as selected one of a set of prestored or loadable reconstruction implementations, with the selection occurring at the beginning of or during a communication.

At block 56, the reconstructed pulse sequence is modified based on the pulse amplitude modification parameters and/or on the pulse position within the frame. For example, the reconstructed pulse sequence can be modified by applying a pulse-amplitude gain scaling function to the subframe whereby the applied gain scaling is a function of position within the subframe. The operations depicted at block 55 are typically duplicated in the decoder and the results or output of the operations depicted at block 55 are passed to a synthesis filter for purposes of reconstructing a version of the original speech. The operations depicted at block 56 are included in the decoder as well and typically, but not necessarily, these operations match the operations of the corresponding pulse sequence modifier of the MPA unit 14. If a pulse-train analysis is utilized, unit 53 becomes a pulse train location determiner and unit 55 becomes a

pulse train reconstruction unit. In addition, the long-term contribution is often utilized in determining the spacing of pulses within a given pulse train.

The operation at blocks 55, 56 and 57 can be implemented as part of the target vector matcher 28 of FIG. 1. Block 59 depicts the encoding operation that corresponds to the encoder  
5 30 of FIG. 1.

In one embodiment of the present invention, the pulse amplitude modifier unit 52 reduces the first pulse's amplitude of every sequence by 12.5 percent and then increases each successive pulse's amplitude by 6.25 percent. For a sequence of six pulses, this results in a pulse amplitude variation of more than 35 percent. Varying the pulse amplitude during the pulse location search  
10 causes the encoder to choose pulse sequence parameters that are different from those the equal amplitude method would choose, and the result is perceptually enhanced speech.

In another embodiment of the present invention, the pulse sequence modifier unit 56 scales each pulse's amplitude as a function of the pulse's position within the frame. This scaling function is a predetermined function of pulse position known both to the encoder's MPA unit and to the corresponding pulse sequence modifier in the decoder. In one implementation of a typical application, this scaling function is an exponentially based function with a negative second derivative, with a total variation across the frame of approximately 10 to 40 percent. In another implementation, this scaling function is an exponentially based function with a negative second derivative, with a total variation across the frame of approximately 20 to 30 percent. In another  
20 implementation, this scaling function is a linear function with a total variation across the frame of approximately 10 to 30 percent.

In another embodiment of the present invention, the pulse sequence modifier unit 56 adds a value to each nonzero pulse amplitude as a function of the pulse's position within the frame. This additive function is a predetermined function of pulse position known both to the encoder's  
25 MPA unit and to the corresponding pulse sequence modifier in the decoder. For a typical application, this scaling function is based on the excitation signal from previous frames and/or on the long-term characteristics of the input speech signal.

In another embodiment of the present invention, the pulse location determiner unit 53 is modified to account for the pulse modification function used in unit 56. The pulse modification function utilized by unit 56, a function of at least the position in the frame, is thus also used to modify the pulse location criterion used by unit 53 in selecting successive pulse locations.

Typically this pulse location criterion is the cross-correlation between the impulse response input and the target vector input, as determined initially by block 50 and as modified with each successive pulse position by unit 53. The cross-correlation is thus scaled by the same amount that unit 56 would scale a pulse in the same position. If it is desired to apply a bias toward a selection of certain pulse positions, unit 53 can be modified using a different function or modified using the inverse of the pulse amplitude modification function used by unit 56.

In other possible embodiments of the present invention, the amplitude modification functions utilized by units 52 and/or 56 are functions of any one or more of the following: (1) the excitation signal from previous frames; (2) the open loop pitch parameters of the present or any previous frame as determined by a long-term pitch analyzer; and (3) the short-term characteristics of the present or any previous frame as determined by the short-term analyzer.

In addition, the pulse location determiner block 53 can be replaced with a pulse train location determiner and the pulse sequence reconstruction block 55 can be replaced with a pulse-train sequence reconstruction unit in order to implement a varying amplitude multi-pulse-train analysis system. The system can then optionally perform both a pulse-sequence analysis and a pulse-train analysis and choose the result that produces a closer match to the target vector.

It will be appreciated that the blocks shown in the above figures can be implemented on a digital signal processing chip, or in software operating on a general purpose processor.

Alternatively, these illustrated embodiments can be implemented using a multi-processor circuit implementation such as described in connection with pending U.S. Patent Application No.

09/005,053, incorporated herein by reference, and such an implementation contemplates the speech data being processed in a circuit that is discrete with respect to a circuit for processing video data as well as a single circuit that processes both the speech and the video data.

Accordingly, the present invention provides a number of advantages. These advantages include, among others, embodiments realizing desirable voice/sound modifications and certain noise reduction qualities. The various embodiments described above are provided by way of illustration only and are not intended to limit the invention. Those skilled in the art will readily recognize various modifications and changes that may be made to the present invention, without strictly following the example embodiments and applications illustrated and described herein. For example, it will be appreciated that blocks 52 and 56 permit many more embodiments than are described here, and that generally the list of preferred embodiments described above is in no way meant to be exhaustive of the set of possible embodiments of this invention. Further, variations on the example operations can be made for a given design specification. Thus, the present invention is not limited by the example embodiments; rather, the scope of the present invention is set forth in the following claims.



What is claimed is:

1. In a speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, a method of analyzing the input speech signal comprising:

generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences having a different average amplitude value; and

outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

2. A system according to claim 1, wherein the target vector is matched using a perceptual weighting criterion.

3. A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

means for generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences having a different average amplitude value; and

means for outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

4. A system according to claim 3, wherein the target vector is matched using a perceptual weighting criterion.

5. A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

an analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable-amplitude pulses, each of said sequences having a different average amplitude value;

the analyzer being further adapted to output a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

6. A system according to claim 5, wherein the target vector is matched using a perceptual weighting criterion.

7. A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

a multi-pulse analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable-amplitude, variable-sign and variably-spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs;

the multi-pulse analyzer being further adapted to output a signal corresponding to a sequence of equal-amplitude, variable-sign, variably-spaced pulses which, according to a maximum likelihood criterion, most closely represents the target vector.

8. A system according to claim 7, wherein the target vector is matched using a perceptual weighting criterion.

9. A system according to claim 7, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the

input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

10. A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a target vector generator for generating data including a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable amplitude, variable sign, variably-spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulses which, according to a maximum likelihood criterion, most closely represents said target vector.

11. A system according to claim 10, wherein the target vector is matched using a perceptual weighting criterion; and

wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

12. A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

6 a multi-pulse analyzer connected to an output line of said target vector generator and an  
7 output line of said short term analyzer, wherein said multi-pulse analyzer generates a plurality of  
8 sequences of variable amplitude, variable sign, variably spaced pulses, each of said sequences  
9 having a different average amplitude value, each of said pulses within each sequence having  
10 variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal  
11 corresponding to the sequence of variable amplitude, variable sign, variably spaced pulses which,  
12 according to the maximum likelihood criterion, most closely represents said target vector.

1 13. A system according to claim 12, wherein the target vector is matched using a perceptual  
2 weighting criterion.

3 14. A system according to claim 13, wherein the pulse amplitude variations are based on at  
4 least one of: the exponential function; a linear function; the short-term characteristics of the  
5 input speech signal; the long-term characteristics of the input speech signal; and the excitation  
6 signal from previous frames.

7 15. A speech processing system comprising:  
8 a short-term analyzer that analyzes an input speech signal, and in response to said input  
9 speech signal, generates the short-term characteristics of the input speech signal;  
10 a target vector generator for generating a target vector from at least said input speech  
11 signal, and optionally, said short-term characteristics; and  
12 a multi-pulse analyzer connected to an output line of said target vector generator and an  
output line of said short term analyzer, wherein said multi-pulse analyzer generates a plurality of  
sequences of variable amplitude, variable sign, variably spaced pulses, each of said sequences  
having a different average amplitude value, each of said pulses within each sequence having  
variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal  
corresponding to the sequence of variable amplitude, variable sign, variably spaced pulses which,  
according to the maximum likelihood criterion, most closely represents said target vector, and

13 one or more pulse sequence modifiers, each having as input at least a sequence of equal  
14 amplitude, variable sign, variably spaced pulses, wherein each said pulse sequence modifier  
15 modifies its input sequence and produces as output a sequence of variable amplitude, variable  
16 sign, variably spaced pulses.

1 16. A system according to claim 15 wherein the pulse sequence modification function is  
2 based on at least one of: the exponential function; a linear function; the short-term  
3 characteristics of the input speech signal; the long-term characteristics of the input speech signal;  
4 and the excitation signal from previous frames.

1 17. A speech processing system comprising:  
2 a short-term analyzer that analyzes an input speech signal, and in response to said input  
3 speech signal, generates the short-term characteristics of the input speech signal;  
4 a long-term analyzer that analyzes an input speech signal, and in response to said input  
5 speech signal, generates the long-term characteristics of the input speech signal;  
6 a target vector generator for generating a target vector from at least said input speech  
7 signal, and optionally, said short-term characteristics, and optionally, said long-term  
8 characteristics; and  
9 a pulse-train sequence analyzer connected to at least an output line of said target vector  
10 generator and an output line of said short term analyzer, wherein said pulse-train sequence  
11 analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced  
12 pulse trains, each of said sequences having a different average amplitude value, each of said  
13 pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train  
14 sequence analyzer for outputting a signal corresponding to the sequence of equal amplitude,  
15 variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion,  
16 most closely represents said target vector.

1 18. A system according to claim 17, wherein the pulse amplitude variations are based on at  
2 least one of: the exponential function; a linear function; the short-term characteristics of the  
3 input speech signal; the long-term characteristics of the input speech signal; and the excitation  
4 signal from previous frames.

1 19. A system according to claim 18, wherein the target vector is matched using a perceptual  
2 weighting criterion.

1 20. A speech processing system comprising:

2 a short-term analyzer that analyzes an input speech signal, and in response to said input  
3 speech signal, generates the short-term characteristics of the input speech signal;

4 a long-term analyzer that analyzes an input speech signal, and in response to said input  
5 speech signal, generates the long-term characteristics of the input speech signal;

6 a target vector generator for generating a target vector from at least said input speech  
7 signal, and optionally, said short-term characteristics, and optionally, said long-term  
8 characteristics; and

9 a pulse-train sequence analyzer connected to at least an output line of said target vector  
10 generator and an output line of said short term analyzer, wherein said pulse-train sequence  
11 analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced  
12 pulse trains, each of said sequences having a different average amplitude value, each of said  
13 pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train  
14 sequence analyzer for outputting a signal corresponding to the sequence of variable amplitude,  
15 variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion,  
16 most closely represents said target vector.

1 21. A system according to claim 20, wherein the target vector is matched using a perceptual  
2 weighting criterion.

1 22. A system according to claim 20, wherein the pulse amplitude variations are based on at  
2 least one of: the exponential function; a linear function; the short-term characteristics of the  
3 input speech signal; the long-term characteristics of the input speech signal; and the excitation  
4 signal from previous frames.

1 23. A system according to claim 21, wherein the pulse amplitude variations are based on at  
2 least one of: the exponential function; a linear function; the short-term characteristics of the  
3 input speech signal; the long-term characteristics of the input speech signal; and the excitation  
4 signal from previous frames.

1 24. A system according to claim 21 wherein the pulse amplitude variations are based on at  
2 least one of: the exponential function; a linear function; and characteristics of the input speech  
3 signal.

1 25. A speech processing system comprising:  
2 a short-term analyzer that analyzes an input speech signal, and in response to said input  
3 speech signal, generates the short-term characteristics of the input speech signal;  
4 a long-term analyzer that analyzes an input speech signal, and in response to said input  
5 speech signal, generates the long-term characteristics of the input speech signal;  
6 a target vector generator for generating a target vector from at least said input speech  
7 signal, and optionally, said short-term characteristics, and optionally, said long-term  
8 characteristics; and  
9 a pulse-train sequence analyzer connected to at least an output line of said target vector  
10 generator and an output line of said short term analyzer, wherein said pulse-train sequence  
11 analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced  
12 pulse trains, each of said sequences having a different average amplitude value, each of said  
13 pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train  
14 sequence analyzer for outputting a signal corresponding to the sequence of variable amplitude,

15 variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion,  
16 most closely represents said target vector, and  
17 one or more pulse-train sequence modifiers, each having as input at least a sequence of  
18 equal amplitude, variable sign, variably spaced pulse trains, wherein each said pulse sequence  
19 modifier modifies its input sequence and produces as output a sequence of variable amplitude,  
20 variable sign, variably spaced pulse trains.

1 26. A system according to claim 25, wherein the target vector is matched using a perceptual  
2 weighting criterion.

1 27. A system according to claim 25, wherein the pulse amplitude variations are based on at  
2 least one of: the exponential function; a linear function; the short-term characteristics of the  
3 input speech signal; the long-term characteristics of the input speech signal; and the excitation  
4 signal from previous frames.

1 28. A system according to claim 25, wherein the pulse-train sequence modification function  
2 is based on the exponential function.

1 29. A system according to claim 25, wherein the pulse-train sequence modification function  
2 is based on a linear function.

1 30. A system according to claim 25, wherein the pulse-train sequence modification function  
2 is based on the short-term characteristics of the input speech signal.

1 31. A system according to claim 25, wherein the pulse-train sequence modification is based  
2 on the long-term characteristics of the input speech signal.



- 1 32. A system according to claim 25, wherein the pulse-train sequence modification function
- 2 is based on the excitation signal from previous frames.

### **Abstract**

A speech signal processing approach modifies the amplitudes of pulses within a multi-pulse sequence to improve and/or modify the perceived quality of reconstructed speech. According to one embodiment that is consistent with the present invention, an input frame processing arrangement generates the short-term characteristics of an input speech signal and also a target vector. The processing arrangement includes an analyzer that operates to provide an optimal analysis, from a maximum-likelihood standpoint, with respect to determining the best possible pulse sequence to match the target. The analyzer receives the target vector and the short term characteristics and generates a plurality of sequences of variable-amplitude pulses, each of said sequences having a different average amplitude value. The analyzer is further adapted to output a signal corresponding to a sequence of either equal-amplitude or unequal-amplitude pulses which, according to a maximum likelihood criterion, would closely represent the target vector.

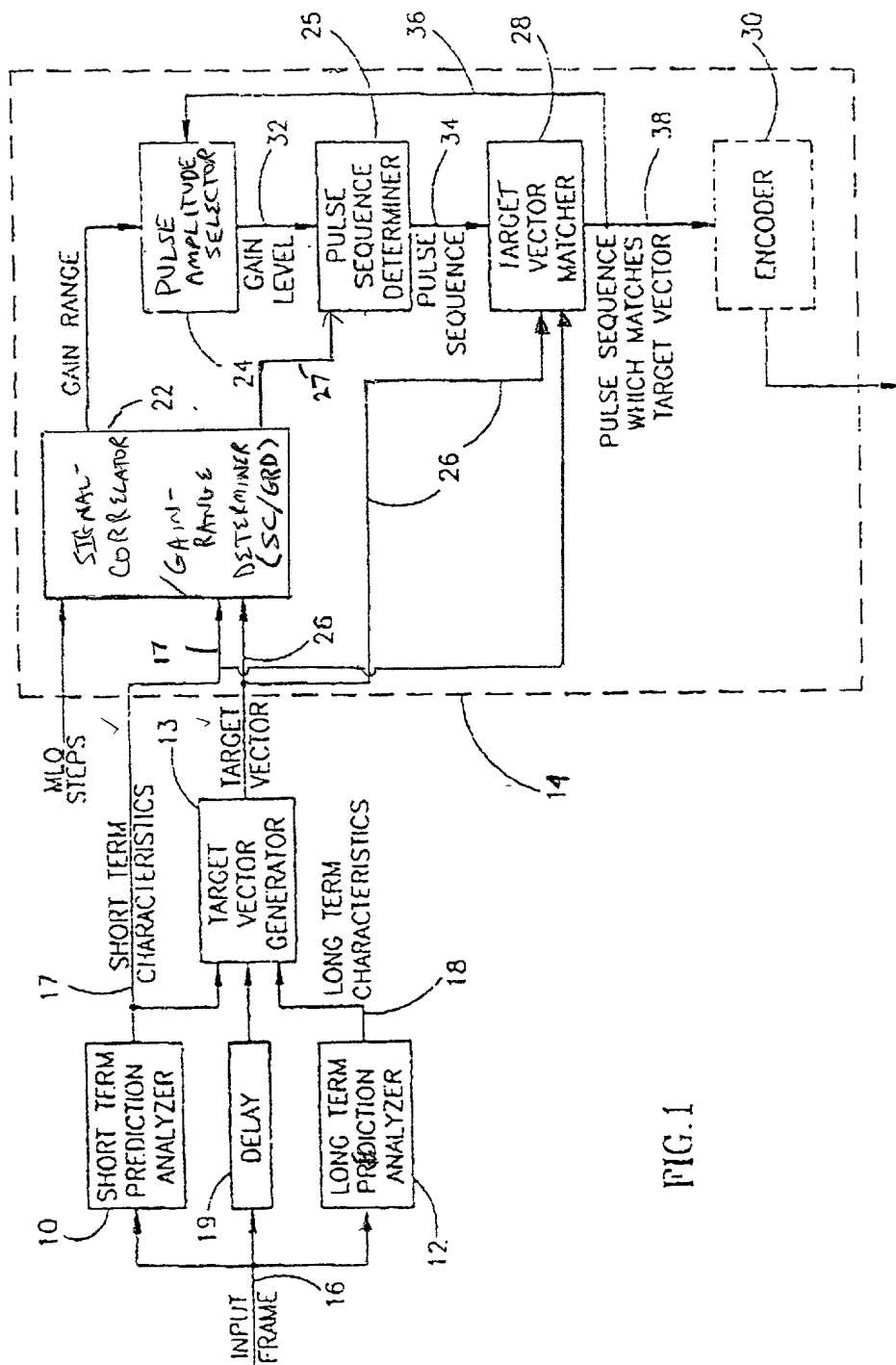


FIG. 1

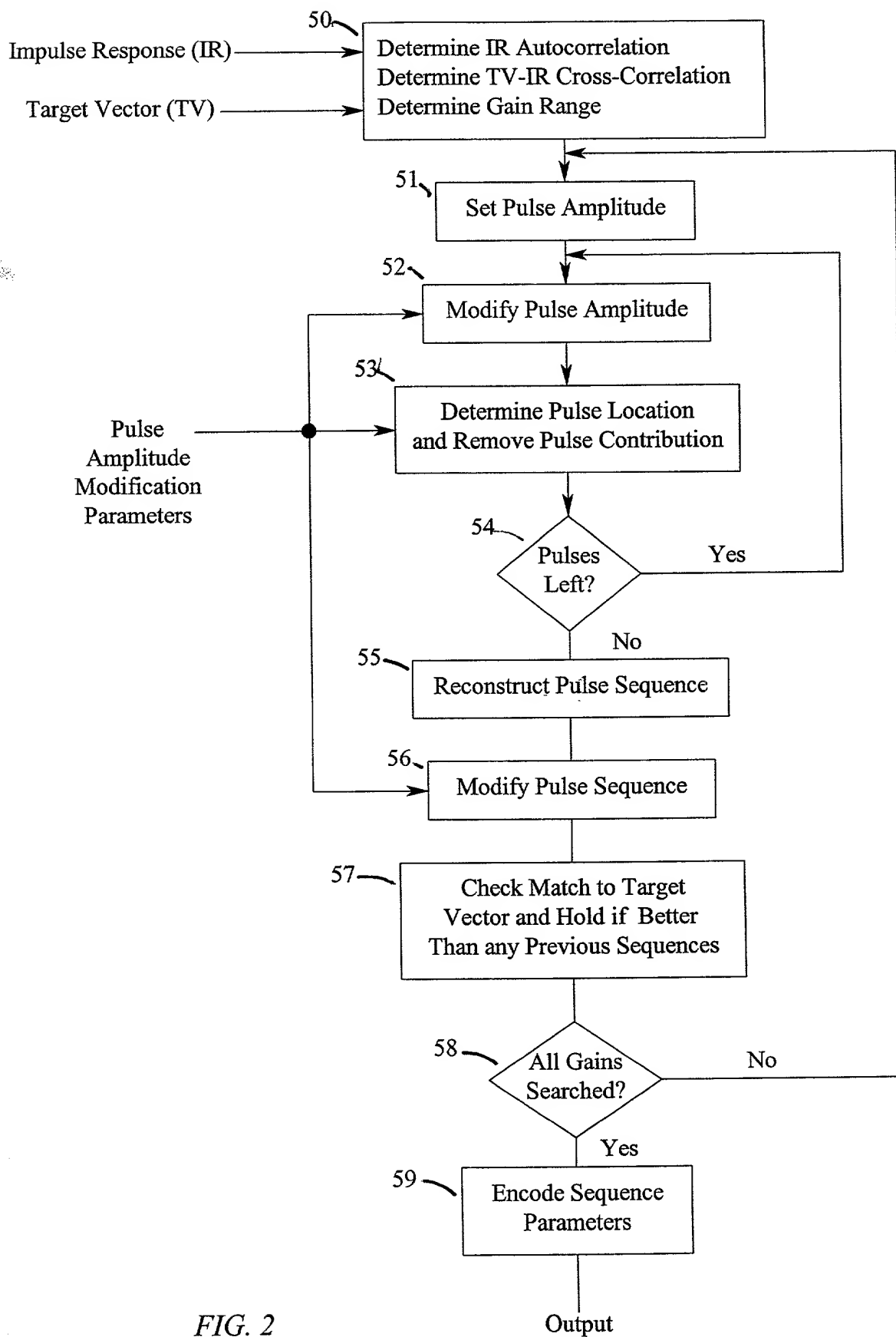


FIG. 2

CRAWFORD PLLC

## United States Patent Application

## DECLARATION UNDER 37 C.F.R. § 1.63

As a below named inventor I hereby declare that: my residence, post office address and citizenship are as stated below next to my name; that

I verily believe I am the original, first and sole inventor (if only one name is listed below) or a joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: **VARYING PULSE AMPLITUDE MULTI-PULSE ANALYSIS SPEECH PROCESSOR AND METHOD.**

The specification of which

- a. ☒ is attached hereto  
 b. ☒ is entitled **VARYING PULSE AMPLITUDE MULTI-PULSE ANALYSIS SPEECH PROCESSOR AND METHOD**, having attorney docket number 8X8S.239PA.  
 c. ☐ was filed on \_\_\_\_\_ as application serial no. \_\_\_\_\_ and was amended on \_\_\_\_\_ (if applicable) (in the case of a PCT-filed application) described and claimed in international no. \_\_\_\_\_ filed \_\_\_\_\_ and as amended on \_\_\_\_\_ (if any), which I have reviewed and for which I solicit a United States patent.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations, § 1.56 (attached hereto).

I hereby claim foreign priority benefits under Title 35, United States Code, § 119/365 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on the basis of which priority is claimed:

- a. ☒ no such applications have been filed.  
 b. ☐ such applications have been filed as follows:

FOREIGN APPLICATION(S), IF ANY, CLAIMING PRIORITY UNDER 35 USC § 119			
COUNTRY	APPLICATION NUMBER	DATE OF FILING (day, month, year)	DATE OF ISSUE (day, month, year)
ALL FOREIGN APPLICATION(S), IF ANY, FILED BEFORE THE PRIORITY APPLICATION(S)			
COUNTRY	APPLICATION NUMBER	DATE OF FILING (day, month, year)	DATE OF ISSUE (day, month, year)

I hereby claim the benefit under Title 35, United States Code, § 120/365 of any United States and PCT international application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. APPLICATION NUMBER	DATE OF FILING (day, month, year)	STATUS (patented, pending, abandoned)

I hereby claim the benefit under Title 35, United States Code § 119(e) of any United States provisional application(s) listed below:

U.S. PROVISIONAL APPLICATION NUMBER	DATE OF FILING (Day, Month, Year)

I hereby authorize them to act and rely on instructions from and communicate directly with the person/assignee/attorney/firm/ organization who/which first sends/sent this case to them and by whom/which I hereby declare that I have consented after full disclosure to be represented unless/until I instruct Crawford PLLC to the contrary.

Please direct all correspondence in this case to Crawford PLLC at the address indicated below:

Crawford PLLC  
333 Washington Avenue North  
Suite 5000  
Minneapolis, MN 55401

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

2	Full Name Of Inventor	Family Name CHRISSAN	First Given Name DOUGLAS	Second Given Name A.
0	Residence & Citizenship	City	State or Foreign Country	Country of Citizenship USA
1	Post Office Address	Post Office Address	City	State & Zip Code/Country
Signature of Inventor 201:				Date:
2	Full Name Of Inventor	Family Name SUBRAMANIAN	First Given Name RAJARATHINAM	Second Given Name G.
0	Residence & Citizenship	City	State or Foreign Country	Country of Citizenship INDIA
2	Post Office Address	Post Office Address	City	State & Zip Code/Country
Signature of Inventor 202:				Date:
2	Full Name Of Inventor	Family Name	First Given Name	Second Given Name
0	Residence & Citizenship	City	State or Foreign Country California	Country of Citizenship U.S.A.
3	Post Office Address	Post Office Address	City	State & Zip Code/Country
Signature of Inventor 203:				Date:
2	Full Name Of Inventor	Family Name	First Given Name	Second Given Name
0	Residence & Citizenship	City	State or Foreign Country	Country of Citizenship
4	Post Office Address	Post Office Address	City	State & Zip Code/Country
Signature of Inventor 204:				Date:

### § 1.56 Duty to disclose information material to patentability.

(a) A patent by its very nature is affected with a public interest. The public interest is best served, and the most effective patent examination occurs when, at the time an application is being examined, the Office is aware of and evaluates the teachings of all information material to patentability. Each individual associated with the filing and prosecution of a patent application has a duty of candor and good faith in dealing with the Office, which includes a duty to disclose to the Office all information known to that individual to be material to patentability as defined in this section. The duty to disclose information exists with respect to each pending claim until the claim is canceled or withdrawn from consideration, or the application becomes abandoned. Information material to the patentability of a claim that is canceled or withdrawn from consideration need not be submitted if the information is not material to the patentability of any claim remaining under consideration in the application. There is no duty to submit information which is not material to the patentability of any existing claim. The duty to disclose all information known to be material to patentability is deemed to be satisfied if all information known to be material to patentability of any claim issued in a patent was cited by the Office or submitted to the Office in the manner prescribed by §§ 1.97(b)-(d) and 1.98. However, no patent will be granted on an application in connection with which fraud on the Office was practiced or attempted or the duty of disclosure was violated through bad faith or intentional misconduct. The Office encourages applicants to carefully examine:

- (1) prior art cited in search reports of a foreign patent office in a counterpart application, and
- (2) the closest information over which individuals associated with the filing or prosecution of a patent application believe any pending claim patentably defines, to make sure that any material information contained therein is disclosed to the Office.

(b) Under this section, information is material to patentability when it is not cumulative to information already of record or being made of record in the application, and

- (1) It establishes, by itself or in combination with other information, a prima facie case of unpatentability of a claim;
- or
- (2) It refutes, or is inconsistent with, a position the applicant takes in:
  - (i) Opposing an argument of unpatentability relied on by the Office, or
  - (ii) Asserting an argument of patentability.

A prima facie case of unpatentability is established when the information compels a conclusion that a claim is unpatentable under the preponderance of evidence, burden-of-proof standard, giving each term in the claim its broadest reasonable construction consistent with the specification, and before any consideration is given to evidence which may be submitted in an attempt to establish a contrary conclusion of patentability.

(c) Individuals associated with the filing or prosecution of a patent application within the meaning of this section are:

- (1) Each inventor named in the application;
- (2) Each attorney or agent who prepares or prosecutes the application; and
- (3) Every other person who is substantively involved in the preparation or prosecution of the application and who is associated with the inventor, with the assignee or with anyone to whom there is an obligation to assign the application.

(d) Individuals other than the attorney, agent or inventor may comply with this section by disclosing information to the attorney, agent, or inventor.